REMARKS:

This paper is herewith filed in response to the Examiner's final Office Action mailed on July 25, 2008 for the above-captioned U.S. Patent Application. This office action is a rejection of claims 1-3, 6-21, 24-39, and 42-46 of the application.

More specifically, the Examiner has rejected claims 1-3, 6-7, 9-17, 19-21, 24-25, 27-35, 37-39, and 42 under 35 USC 103(a) as being unpatentable over Friedman (US6,182,029) in view of Brecher (US7,054,754) and in view of Moore (US5,577,239) in view of Dittmar (J. Chem. Inf. Comput. Sci. Vol. 23, No. 3), in view of Hull (US6,332,138) and in view of Leiter (J. Chem. Doc., Vol. 15, 1965); rejected claims 8 and 26 under 35 USC 103(a) as being unpatentable over Friedman in view of Brecher, in view of Moore, in view of Dittmar, in view of Hull (US6,332,138), and in view of Leiter, and further in view of Drefahl (J. Chem. Inf. Comput. Sci., Vol. 33) and Murray-Rust (New Chem. Vol. 25, p 618-634, 2001); rejected claims 18 and 36 under 35 USC 103(a) as being unpatentable over Friedman in view of Brecher in view of Moore in view of Dittmar, in view of Hull, in view of Leiter, and further in view of Kemp (Chem. Inf. Comput. Sci., Vol. 38); rejected claims 43-44 under 35 USC 103(a) as being unpatentable over Friedman in view of Brecher and in view of Moore in view of Dittmar in view of Hull in view of Leiter, and in view of Shivaratri (Computer December 1992); and rejected claims 45-46 under 35 USC 103(a) as being unpatentable over Friedman in view of Brecher and in view of Moore in view of Dittmar in view of Hull in view of Leiter in view of Shivaratri in view of Drefahl and Murray-Rust. The Applicants disagree with the rejections.

Claims 1, 3, 7, 19, 21, 37-39, and 43 have been amended merely to improve their clarity and to correct typographical errors. No new matter is added.

The Applicants submit that the amendments are not seen to be substantive, and the amendments do not require a further search by the Examiner. Thus, the Examiner is respectfully requested to enter these amendments in this Response to final Office Action.

Regarding the rejection of claim 1 under 35 USC 103(a) the Examiner states:

> "Friedman shows a method and system for extracting information from natural language text data. Friedman shows information is extracted from text documents (col. 4, line 59-63). Friedman shows that the text of the text document is partitioned into phrases (col. 6, line 36-45). **Friedman shows that partitioned phrases are further parsed to assign semantic meaning to words** (col. 6, line 63-65). **Friedman suqqests that chemical information can be identified and extracted** (col. 11, line 34-50). Friedman shows the method provides reliable and efficient access to information within a document and is useful for retrieving and summarizing relevant information in documents (col. 4, line 59-67)," (emphasis added).

Firstly, the Applicants submit that the Examiner is in error where the Examiner appears to assert that Friedman relates to **recognizing chemical name fragments** as in claim 1.

Friedman discloses:

> "The preprocessor 11 thus performs lexical lookup to **identify and categorize multi-word and single word phases within each sentence**. The output of this component consists of a list of word positions where each position is associated with a word or multi-word phrase in the report. For example, assuming that the sentence "spleen appears to be moderately enlarged" is at the beginning of the report, it would be represented as the list where position 1 is associated with "spleen", position 2 with the multi-word phrase "appears to be", position 5 with "moderately", and 6 with "enlarged". The remainder of the list of word positions would be associated with the remaining words in the report," (emphasis added), (col. 7, lines 17-29).

The Applicants submit that Friedman appears to **identify and categorize all words and phrases in a sentence**.

Friedman discloses:

> "The information extraction system of FIG. 1, known as **MedLEE, is designed for use as a general processor within the medical domain, e.g., radiography, mammography, neuroradiology, pathology, and electrocardiography**, etc. Although used for language extraction in the medical/[clinical] context, MedLEE

can be adapted for use in other domains such as genomics." (col. 6, lines 12-19).

The Applicants submit that although language extraction may include radiography, mammography, neuroradiology, pathology, and electrocardiography, as stated above, there can not be found in all of Friedman any support for a method or an apparatus which discloses or suggests "partitioning text of the text document and assigning semantic meaning to words of the partitioned text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries **to recognize chemical name fragments**," as in claim 1.

In the rejection the Examiner appears to argue that Friedman relates to **recognizing chemical names** in a text document, since the Examiner states that "Friedman suggests that chemical information can be identified and extracted, (col. 11, line 34-50)."

As cited Friedman discloses:

> "FIG. 3 shows a block diagram of a second embodiment of the information extraction (MedLEE) program of FIG. 1. The modified program 300 includes **a tagger routine 16 for linking the structured output described previously with respect to FIG. 1** to the corresponding words in the original sentences of the text data input. **Preferably, the tagger 16 utilizes markup languages**, such as Hypertext Markup Language (HTML) and Extensible Markup Language (XML), which are derived from Standard Generalized Markup Language (SGML) and which are used rendering documents for the World Wide Web. **Widespread adoption of markup languages are evidenced by**: the Text Encoding Initiative (TEI) which uses SGML to encode literature; **Chemical Markup Language (CML), which involves documentation of chemical compounds using SGML**; and Open Financial Exchange (OFE), which is an SGML standard format for interchange of financial transactions," (col. 11, lines 34-50)," (emphasis added).

The Applicants contend that the Examiner has not shown where, as stated in the rejection, "Friedman suggests that chemical information can be identified and extracted." The Applicants submit that here, Friedman discloses a tagger routine for linking a structured output to corresponding words and that the tagger preferably utilizes markup languages. The Applicants submit that as cited Friedman is merely seen to provide a comment that there is "**Widespread**

adoption of markup languages are evidenced by: the Text Encoding Initiative (TEI) which uses SGML to encode literature; **Chemical Markup Language (CML), which involves documentation of chemical compounds using SGML**; and Open Financial Exchange (OFE), which is an SGML standard format for interchange of financial transactions." The Applicants contend that these apparent examples of a "Widespread adoption of markup languages" can not be seen to support the assertion by the Examiner in the rejection that "Friedman suggests that chemical information can be identified and extracted."

Further in the rejection of claim 1 the Examiner states:

> **Friedman does not show the application of regular expressions and a plurality of chemical dictionaries to recognize chemical names** or storing information in a searchable index.
>
> "Brecher shows a method system and computer program product for **processing text documents to extract chemical information**. Brecher shows the application of regular expression (col. 5, line 41-45) and a plurality of dictionaries to recognize chemical names (col. 6, line 29-40). Brecher shows that the lexicon has at least a sub lexicon to identify stopwords (col. 8, line 49-50), prefixes (001. 9, line 55) or suffixes (col. 11, line 43). Brecher shows that substructures are recognized (col. 6, line 31-33). Brecher shows that structural connectivity is determined (col.7, line 35-57). Brecher et al. shows that identifying information is extracted from the substructures and fragments to produce a fully parsed chemical name that is correlated to a chemical structure. Brecher shows the method allows chemical names to be accurately converted to chemical structures in real time or in nearly real time to provide users with a powerful, practical tool (col. 2, line 11-14)," (emphasis added).

Firstly, the Applicants note that here the Examiner appears to contradict his previous assertion in the rejection that "Friedman suggests that **chemical information can be identified and extracted**, (col. 11, line 34-50)," (emphasis added). The Applicants submit that the rejection is unclear as to whether it is Friedman or Brecher who is allegedly performing a function of identifying and extracting chemical information as asserted by the Examiner. The Applicants contend that neither Friedman nor Brecher can be seen to disclose or suggest "**partitioning text of the text document** and assigning semantic meaning to words of the partitioned text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of

dictionaries **to recognize chemical name fragments**," as in claim 1.

Regarding the Examiner's statement in the rejection that "Brecher shows a method system and computer program product for **processing text documents to extract chemical information**," the Applicants note that Brecher discloses:

> "**A chemical name 12 is supplied via one or more input systems such as end-user keyboard input 14, file-based input 16, or World-Wide Web query input 18**," (emphasis added), (col. 2, lines 47-50).

That is, Brecher discloses **a chemical name is supplied via a keyboard, file-base input, or a query input**. The Applicants can not find in all of Brecher where Brecher discloses **partitioning text of the text document to recognize chemical name fragments**. Further, the Applicants note that as stated above the Examiner admits that "Friedman does not show the application of regular expressions and a plurality of chemical dictionaries to recognize chemical names."

The Applicants contend that a person skilled in the art would not be motivated to combine Friedman and Brecher, though not agreed to as proper, for at least the reason that such a combination would still not disclose or suggest at least where claim 1 recites "**partitioning text of the text document** and assigning semantic meaning to words of the partitioned text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries **to recognize chemical name fragments**."

Further, the Applicants note that in addition to Friedman and Brecher, the Examiner has generally combined the references Dittmar, Hull, Leiter, and Moore in order to reject where claim 1 recites in part:

> "extracting information associated with the recognized chemical name fragments and substructures of the text document and indexing the extracted information in a text index;
>
> indexing representations of the recognized chemical name fragments and the

> substructures in association with the determined structural connectivity information into a plurality of chemical connectivity tables;
>
> storing the text index in association with the indexed representations in a searchable index; and
>
> providing a graphical user interface to search the index, where the search comprises entering one or more chemical fragment names and entering one or more substructures in the representation form, where the entering is by at least one of text form or graphical selection."

The Applicants submit that the rejection is unclear for at least the reason that the Examiner has not applied these references to individual elements of claim 1. The Examiner appears to argue that the references Dittmar, Hull, Leiter, and Moore combined would somehow disclose or suggest the remaining elements of claim 1 as stated above. The Applicants strongly disagree with the rejection and request clarification in a non-final Office Action.

Dittmar relates to using a CAS online system to search chemical structures and nomenclature data, (page 93, col. 1, introduction).

In the Office Action, the Examiner states:

> "Dittmar et al. shows searching an index by at least one of a fragment or substructure connectivity (p.99, col. 2, para2, sent. 1) using a graphical user interface (p. 93, col. 1, para. 3, sent. 2). **Dittmar et al. shows characters comprising at least one of upper case C, 0, R, N, H** (p. 98, col. 1 par 2, sent. 3; para.3, sent. 1; and p. 99, para 2-3). Dittmar et al. shows the implementation of a user interface to simplify searching (p. 93, col. 1, para 3, sent. 1)," (emphasis added).

The Applicants respectfully point out to the Examiner that Brecher, which the Examiner appears to combine with Dittmar, discloses:

> "In an initial stage of the preprocessing, the individual characters of the name are manipulated as follows without reference to the chemical meaning implied by the characters. **The name is converted to all lower-case characters**," (emphasis

added), (col. 3, lines 4-8).

The Applicants submit that the proposed combination, though not agreed with, is improper at least for the reason that the Examiner appears to be asserting a feature, in the rejection, of searching upper case characters in Dittmar which appears to conflict with a name being converted to all lower-case characters as in Brecher.

In addition, the Applicants submit that Dittmar appears to utilize a search method to search a specifically structured proprietary database of the Chemical Abstract Service (CAS) Chemical Registry System (see page 1, col. 1, lines 1-4). The Applicants submit that one skilled in the art would not be motivated to combine a user interface of Dittmar to search other than CAS systems.

Further, the Applicants submit that the specific application of Hull in the rejection is unclear. It appears that in the rejection the Examiner applies Hull to extracting keywords, storing identifying information and keywords in an index, and searching the index.

The Applicants note that Hull relates to a method for computing chemical similarities using textual and chemical descriptors, (abstract). Hull discloses that two databases are generated for a comparison and that these data bases are called $TIMI_T$ and $TIMI_C$. Further, Hull appears to indicate that these two data bases are combined in another database called $TIMI_{TC}$. In addition, Hull discloses that the "T" stands for text and "C" stands for chemistry, (col. 12, lines 44-53).

In the Office Action the Examiner states:

> "**Hull et al shows extracting keywords and storing identifying information and keywords in association with structural connectivity in a searchable index and then searching the index using a keyword and a chemical fragment**. Hull teaches extracting keywords from the document (col. 9, lines 15 - 32). Extracted identifying information is stored in association with structural connectivity information in a searchable matrix (index) (col.10, lines 32-52). **Hull et al. teach the searching of the index by a keyword and a fragment/substructure name or connectivity** (col. 16, lines 21-33, and col. 13, lines 40-67)," (emphasis added).

As cited Hull discloses:

> "**Generating a TIMI database includes** the following sequential, non-sequential, or sequence independent steps. Referring to FIG. 3, in step S300, **a user and/or a computer generates or creates chemical and textual descriptors for each compound represented in the database**. The textual descriptors may, for example, originate from a collection of documents, or other text source, in, say, ASCII format or other suitable format. A textual representation of the chemical descriptors is also added to the textual descriptors," (col. 9, lines 15-25), (emphasis added).

The Applicants submit that, as cited, Hull can not be seen to be **extracting keywords** as indicated in the rejection. Here, Hull discloses a method for generating a TIMI database by **generating descriptors for each compound represented in a database**. The Applicants submit that the Examiner's application of Hull in the rejection is unclear. Further, the Applicants contend that Hull can not be seen to address a shortfall of Brecher or Friedman as stated above.

Further, as cited Hull discloses:

> "Finally, we can perform one special search in $TIMI_{TC}$ that can not be performed in either $TIMI_T$ or $TIMI_C$ individually--a combined structure and text query. Combining both query types is advantageous because one can "tweak" a structural search with carefully chosen keywords. For example, suppose the user is most interested in the possibility of toxicity with a given compound. **She can then add terms related to toxicity to the structural query, thereby ranking documents which discuss toxicity issues more highly**," (emphasis added), (col. 16, lines 23-32).

It is noted that here Hull discloses entering terms to a structural query. The Applicants contend that Hull can not be seen to disclose or suggest at least where claim 1 recites "to search the searchable index, where the search comprises **entering one or more chemical fragment names and entering one or more substructures in a representation form**, where the entering is by at least one of text form or graphical selection."

Further, the Applicants note that, similar to Dittmar, Hull also appears to perform a search adapted for specifically structured proprietary databases (i.e. TIMI databases). Hull discloses that the TIMI databases are generated according to specific procedures (see col. 9, line 15 to col. 10, line 52). The Applicants submit that one skilled in the art would not be motivated to combine the references for at least the reason that the references appear to rely on different search methodologies based on disparate proprietary database systems.

Further, in the rejection the Examiner appears to apply Moore for storing information in a searchable index, and searching the index. In the rejection the Examiner states:

> "Moore et al. shows a method of storing extracted identifying information in a searchable index (col. 4, line 28-35). **Moore et al. shows that the index can be searched by a combination of substructure names, reading on text terms and keywords** (col. 7, line 47-48) and connectivities, reading on graphical representations (col. 10, line 43-46). **Moore shows that multiple databases can be interrelated to form a relational database forming an integrated chemical information system that can be searched advantage of simplified search queries** (col. 12, line 42-46). Moore et al. shows the method has the further advantages of reducing database development and maintenance costs, simplify interfacing with other information systems (col. 2, line 10-23)," (emphasis added).

As cited Moore discloses:

> "V. PROCESSING OF QUERY SUBSTRUCTURE SEARCHES," (col. 7, lines 47-48); and
>
> "Chemical name searching can be accomplished by storing and indexing carefully defined name fragments, as well as indexing the complex strings of the complete chemical names. **Searching can be performed on a partial or complete chemical name query using standard relational database technology**," (emphasis added), (col. 10, lines 42-46).

It is noted that the application of Moore in the rejection is also unclear. There does not appear to be a specific application of the Moore citations in relation to the claims. The Applicants submit that as cited Moore appears to perform a search of a partial or complete chemical name using relational database technology. Further, the Applicants note that in Moore there are also disclosed

dynamic queries and structure class searches. The dynamic queries appear to allow a user to store a search result as dynamic such that a user will be notified if a new substructure, chemical name, or molecular formula is added which matches the earlier stored dynamic search, (col. 11, lines 46-63). Further, as a method to reduce overhead Moore discloses that a structure class search can be used to limit the scope of searches to a class of chemical types (col. 12, lines 37-39).

The Applicants contend that none of these methods in Moore can be seen to relate to a search comprising **entering one or more chemical fragment names and entering one or more substructures in a representation form**, where the entering is by at least one of text form or graphical selection as in claim 1.

Further, the Applicants submit that the reference Leiter is not seen to address the issues as stated above.

For at least these reasons the Applicants contend that the references cited, alone or combined, can not be seen to disclose or suggest claim 1. Thus, the Applicants respectfully request that the Examiner remove the rejection of claim 1 and allow claim 1 to issue.

Further, as the independent claims 19 and 37 distinguish over the references for reasons similar to those detailed for claim 1 as stated above, the references are not seen to disclose or suggest all of claims 1, 19, and 37. Thus, the rejections of these claims should be removed.

It is noted that in the rejection of independent claim 43 the Examiner has cited the Shivaratri reference in addition to the references cited against independent claim 1. Further, the Applicants note that independent claims 43 distinguishes over the references for reasons similar to those detailed for claim 1, and that Shivaratri is not seen to overcome the arguments presented against these references. Thus, the rejection of claim 43 is seen to be overcome and the rejection should be removed.

Further, the Applicants contend that even if the references were combined, which is not agreed to

as proper, for at least the reasons stated the combination would still not disclose or suggest the claims. Thus, the rejection of these claims should be removed.

Regarding the rejections of dependent claims 2-3, 6-7, 9-17, claims 20-21, 24-25, 27-35, and claims 38-39, and 42 the Applicants note that the Examiner has not specifically addressed these rejections.

Claim 2 recites in part:

> "wherein the extracting further comprises extracting keywords from the text document and indexing the keywords in the text index, and wherein the search comprises **selecting a graphical representation of one or more substructures and additionally entering at least one keyword**"

As similarly stated above, the Examiner is unclear as to how the references allegedly disclose or suggest claim 2. However, the Applicants note that in the Office Action the Examiner states:

"Dittmar et al. shows searching an index by at least one of a fragment or substructure connectivity (p.99, col. 2, para2, sent.1*)* using a graphical user interface (p. 93, col. 1, para. 3, sent. 2),"

"Hull et al. teach the searching of the index by a keyword and fragment/substructure name or connectivity (col. 16, lines 21-33, and col. 13, lines 40-67," and

"Leiter et al. shows searching indices to identify documents related to a chemical compound (p. 238, col. 2, lines 5-7)."

Firstly, the Applicants can not find in all of the references where there is disclosed or suggested **selecting a graphical representation** of one or more substructures and entering a keyword. Further, as stated above, Dittmar and Hull are seen to be using disparate database systems and searching methods. For at least these reasons the proposed combination of these references, as in the rejection, is not seen to be proper.

The Applicants contend that the references cited can not be seen to disclose or suggest at least

where claim 2 relates to **selecting a graphical representation of one or more substructures and additionally entering a keyword**. Thus, the rejection of claim 2 should be removed.

In addition, claims 20 and 38 are seen to distinguish over the references for reasons similar to those detailed for claim 2 as stated above. Thus, the rejections of these claims should be removed

Regarding claim 3 the Examiner appears to support the rejection where the Examiner states:

> Hull et al shows extracting keywords and storing identifying information and keywords in association with structural connectivity in a searchable index and then searching the index using a keyword and a chemical fragment. Hull teaches extracting keywords from the document (col. 9, lines 15 -32). Extracted identifying information is stored in association with structural connectivity information in a searchable matrix (index) (col.10, lines 32-52). Hull et al. teach the searching of the index by a keyword and a fragment/substructure name or connectivity (col. 16, lines 21-33, and col. 13, lines 40-67.

The Applicants note that Hull discloses:

> "searching a TIMI database is carried out as follows. In step S400, the **user specifies one or more words and/or chemical structures as a probe**. The connection table of a probe molecule, text, or multiple molecules or text in the case of a joint probe, is converted to the descriptor set of the TIMI database to create a feature, or column, vector for the probe in step S410," (emphasis added),(col. 10, lines 55-61)

The Applicants contend that Hull relates to searching the TIMI database using one or more words and/or chemical structures as a probe. Hull appears to disclose that in response to the search query a connection table of the probe is converted to a descriptor set to create a feature, or column, vector for the probe. Thus, it appears that in Hull there the connection table is utilized only as a result of specifying **words and/or chemical structures**. The Applicant submits that nowhere in Hull can there be found a disclosure or suggestion of at least where claim 3 recites "wherein the search comprises additionally entering at least one keyword, **and at least one of fragment connectivity and substructure connectivity**." For at least this reason the rejection of claim 3 should be removed.

In addition, claims 21 and 39 are seen to distinguish over the reference for reasons similar to those detailed for claim 3 as stated above. Thus, the rejections of these claims should be removed

Regarding the rejection of claim 16 the Applicants note that claim 16 recites:

> A method as in claim 14, where the characters comprise at least one of **upper case C, O, R, N and H.**

The Examiners appears to support the rejection of claim 16 where the Examiner states:

> **"Dittmar et al. shows characters comprising at least one of upper case C, 0, R, N, H** (p. 98, col. 1 par 2, sent. 3)," (emphasis added).

The Applicants submit that, as stated above, Brecher, which the Examiner appears to combine with Dittmar, discloses:

> "**In an initial stage** of the preprocessing, the individual characters of the name are manipulated as follows without reference to the chemical meaning implied by the characters. **The name is converted to all lower-case characters**," (emphasis added), (col. 3, lines 4-8).

Thus, the Applicants submit that as Brecher discloses the initial stage converts a name to all lower-case characters, the application of at least Brecher in view of Dittmar to reject claim 16 can not be seen to be proper. For at least this reason the rejection of claim 16 is seen to be moot and the rejection should be removed.

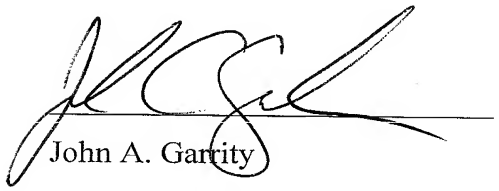
In addition, claim 34 is seen to distinguish over the references for reasons similar to those detailed for claim 16 as stated above. Thus, the rejections of claims 16 and 34 should be removed

In addition, for at least the reason that claims 2-3, and 6-18; claims 20-21, and 24-36; and claims 38-40, and 42; and claims 44-46 depend from independent claims 1, 19, 37, and 43, respectively, the references cited are not seen to disclose or suggest all these claims and the rejection of all these claims should be removed.

Based on the above explanations and arguments, it is clear that the references cited cannot be seen to disclose or suggest claims 1-3, 6-21, 24-40, and 42-46. The Examiner is respectfully requested to reconsider and remove the rejections of all claims and to allow all of the pending claims 1-3, 6-21, 24-40, and 42-46.

For all of the foregoing reasons, it is respectfully submitted that all of the claims now present in the application are clearly novel and patentable over the prior art of record. Should any unresolved issue remain, the Examiner is invited to call Applicants' attorney at the telephone number indicated below.

Respectfully submitted:



9/10/2006

John A. Garrity
Reg. No.: 60,470
Customer No.: 48237
HARRINGTON & SMITH, PC
4 Research Drive
Shelton, CT 06484-6212
Telephone: (203)925-9400
Facsimile: (203)944-0245
email: jgarrity@hspatent.com

Date